

PREAMBLE

Jason is currently studying Scientific Computation and Applied Mathematics. Having switched from theoretical physics he's been teaching himself computer science, probabilistic programming and machine learning. He's usually hacking away at python for the most part and when he need to he can use java just fine.

He spent the first half of this year really understanding how python works. From garbage collection to now python data structures are implemented and how they should be used. After some time with friends, longboarding, ping pong, etc. The rest of this time is spent learning more about Markov chains, bandit algorithms and language processing.

Jason Liu

github//dribbble : jxnl
jason@jxn1.co
6479888953

CORE COMPETENCIES

I'm quite familiar with most of the data related tools in Python such as numpy, pandas, patsy, ggplot and matplotlib. I'm getting better at Java every day and really being to appreciate OOP designs patterns. I try not to confuse my passion with my knowledge but I think I've got a hand on machine learning theory and how to plan out a data science job. For tools, I am extremely comfortable with vim as a usually work on a dev server. I can also pull out some sick graphic design and motion animation if it is ever needed

INTERESTING COURSE WORK

2B Honours Applied Mathematics co-op program.

Grades good enough to be put on a list.

Statistics, Probaility, Computation Physics, Quantum Mechanics.

Online Courses through Coursera

Free education during my free time.*

Machine Learning with Andrew Ng, Data Science with Bill Howe, Applied Math with Nathan Kutz, and Natrual Language Processing with Dan Jurafsky.

EXPERIMENTS & EXPERIENCE

Sysomos — Toronto

Data Scientist Intern

Implimented clustering over twitter subgraphs and community detection over multiple resoltions . Writing map reduce jobs for hdfs to process twitter hashtags graph edges onto TitanDB

PennAppsX — Philly

Qualified in Top 20 out of 300+ projects at PennApps Fall 2014

Awarded Best use of Machine Learning

Implemented machine learning OVA text classification along with spam detection to automate the process of automating facebook groups

Used sklearn and tf-idf vectorizer to train a SVM on facebook group content for topic classification"

Reservoir Sampling for Python

Almost didn't get a techinical question right about reservoir sampling.

Build a python package + cli for streaming data and iterables. Its on PyPi!

Supports uniform, exponential, and weighted sampling with a reservoir

Sentiment Analysis

I am not not happy

Designed a Naive Bayes sentiment classifier that would handle negations.

Visualized various project Gutenberg books and how sentiment changed over time.

Communitation and Outreac — TRIUMF

Paid to be really excited about physics

Organized weekly talks on current research and how it affects the community.